RESEARCH ARTICLE

A convolutional neural network for steady state visual evoked potential classification under ambulatory environment

No-Sang Kwak[1], Klaus-Robert Müller[1,2], Seong-Whan Lee[1]*

1 Department of Brain and Cognitive Engineering, Korea University, Anam-dong, Seongbuk-ku, Seoul, Republic of Korea, **2** Department of Computer Science, TU Berlin, Berlin, Germany

* sw.lee@korea.ac.kr

Abstract

The robust analysis of neural signals is a challenging problem. Here, we contribute a convolutional neural network (CNN) for the robust classification of a steady-state visual evoked potentials (SSVEPs) paradigm. We measure electroencephalogram (EEG)-based SSVEPs for a brain-controlled exoskeleton under ambulatory conditions in which numerous artifacts may deteriorate decoding. The proposed CNN is shown to achieve reliable performance under these challenging conditions. To validate the proposed method, we have acquired an SSVEP dataset under two conditions: 1) a static environment, in a standing position while fixated into a lower-limb exoskeleton and 2) an ambulatory environment, walking along a test course wearing the exoskeleton (here, artifacts are most challenging). The proposed CNN is compared to a standard neural network and other state-of-the-art methods for SSVEP decoding (i.e., a canonical correlation analysis (CCA)-based classifier, a multivariate synchronization index (MSI), a CCA combined with k-nearest neighbors (CCA-KNN) classifier) in an offline analysis. We found highly encouraging SSVEP decoding results for the CNN architecture, surpassing those of other methods with classification rates of 99.28% and 94.03% in the static and ambulatory conditions, respectively. A subsequent analysis inspects the representation found by the CNN at each layer and can thus contribute to a better understanding of the CNN's robust, accurate decoding abilities.

Introduction

A brain-computer interface (BCI) allows for the decoding of (a limited set of) user intentions employing only brain signals—without making use of peripheral nerve activity or muscles [1, 2]. BCIs allow users to harness brain states for driving devices such as spelling interfaces [3–5], wheelchairs [6, 7], computer games [8, 9] or other assistive devices [10–12]. Recent BCI studies have demonstrated the possibility of decoding the user's intentions within a virtual reality environment [13] and using an exoskeleton [14–17]. Others have investigated the decoding of expressive human movement from brain signals [18]. Furthermore, researchers have developed several applications of BCI systems for the rehabilitation of stroke patients [19–23].

Existing EEG-based BCI techniques generally use EEG paradigms such as the modulation of sensorimotor rhythms through motor imagery (MI) [24–26]; event related potentials (ERPs), including P300 [27, 28]; and steady-state visual-, auditory-, or somatosensory-evoked potentials (SSVEPs [29, 30], SSAEPs [31, 32], and SSSEPs [33, 34]).

Of these EEG paradigms, SSVEPs have shown reliable performance in terms of accuracy and response time, even with a small number of EEG channels, at a relatively high information transfer rate (ITR) [35] and reasonable signal-to-noise ratio (SNR) [36]. SSVEPs are periodic responses elicited by the repetitive fast presentation of visual stimuli; they typically operate at frequencies between 1 and 100 Hz and can be distinguished by their characteristic composition of harmonic frequencies [33, 37].

Various machine learning methods are used to detect SSVEPs: first and foremost, classifiers based on canonical correlation analysis (CCA), a multivariate statistical method for exploring the relationships between two sets of variables, can harvest the harmonic frequency composition of SSVEPs. CCA detects SSVEPs by finding the weight vectors that maximize the correlations between the two datasets. In our SSVEP paradigm, the maximum correlation extracted by CCA is used to detect the respective frequencies of the visual stimuli to which the subject attended [37]. Modified CCA-based classifiers have been introduced, such as a multiway extension of CCA [38], phase-constrained [39] and multiset [40] CCA methods. In addition, stimulus-locked intertrace correlation (SLIC) [41] and the sparsity-inducing LASSO-based method [42] have been proposed for SSVEP classification. The multivariate synchronization index (MSI) was introduced to estimate the synchronization between two signals as an index for decoding stimulus frequency [43, 44]. SSVEP decoding can be further extended by employing characteristics based on phase and harmonics [35], boosting the ITRs significantly. Recently, deep-learning-based SSVEP classification methods [45–47] have also been considered; however, all have thus far used prestructuring by employing a Fourier transform in the CNN layer.

Recently, brain machine interface (BMI) researchers have turned their focus to connecting SSVEP-based BMIs to mobile systems with wireless EEG telemetry [14, 48] in order to explore the feasibility of implementing online SSVEP-based BMIs. This progress has greatly facilitated the transition of laboratory-oriented BMI systems to more practical ambulatory brain-controlled exoskeletons [14].

Despite the technical and machine learning progress outlined, systematic performance deterioration between ambulatory and static BMI-control conditions has been found [49, 50], primarily because of the artifacts, which are caused by subject's motion, head swing, walking speed or sound, and the exoskeleton's electric motors [51, 52]; these artifacts may, in addition, differ across users [14].

We address this key challenge by exploring deep learning methods as a means to reliably minimize the influence of artifacts on ambulatory SSVEP-BMIs. Here, we consider as artifacts all signals that have non-cerebral origin and that might mimic non-task-related cognitive signals or that are induced by external factors (e.g., while walking, a subject's head may be moved by the exoskeleton which can give rise to swinging movements in the line between the electrodes and EEG amplifiers, leading to disconnections or high impedance measurements in extreme cases). These artifacts typically distort the analysis of an EEG.

Therefore, we propose a CNN-based classifier that uses frequency features as input for robust SSVEP detection in ambulatory conditions. In the course of the CNN training process, the model learns an appropriate representation for solving the problem [53, 54]. The receptive field/convolution kernel structure of the trained model can then be analyzed, and we can interpret the high-level features found by the deep network as we inspect each layer. Our CNN architecture compares favorably with standard neural network and other state-of-the-art methods used in ambulatory SSVEP BMIs.

Materials and methods

Experiment

Experimental environment. We designed an experimental environment for SSVEP-based exoskeleton control following [14]. In particular, we used a powered lower-limb exoskeleton (Rex, Rex Bionics Ltd.) with a visual stimulus generator attached to the robot for stimulating SSVEPs in an ambulatory environment. The visual stimulus generator presented visual stimuli using five light-emitting diodes (LEDs: 9, 11, 13, 15, and 17 Hz with a 0.5 duty ratio) which were controlled by a micro controller unit (MCU; Atmega128), as shown in Fig 1.

The EEG was acquired from a wireless interface (MOVE system, Brain Products GmbH) using 8 Ag/AgCl electrodes at locations PO7, PO3, PO, PO4, PO8, O1, Oz, and O2, with reference (FCz) and ground (Fpz) electordes, illustrated in Fig 2. Impedances were maintained below 10 kΩ and the sampling frequency rate was 1 kHz. A 60 Hz notch filter was applied to the EEG data for removing AC power supply noise.

Subject. Seven healthy subjects, with normal or corrected-to-normal vision and no history of neurological disease, participated in this study (age range: 24–30 years; 5 males, 2 females). The experiments were conducted in accordance with the principles expressed in the Declaration of Helsinki. This study was reviewed and approved by the Institutional Review Board at Korea University [1040548-KU-IRB-14–166-A-2] and written informed consent was obtained from all participants before the experiments.

Experiment tasks. We acquired two SSVEP datasets under static and ambulatory conditions, respectively, to compare the performance of the SSVEP classifiers. From Task 1, we collected SSVEP data with the exoskeleton in a standing position (static SSVEP). In Task 2 (ambulatory SSVEP), the SSVEP signals were acquired while the exoskeleton was walking. In both tasks, we performed the experimental procedure described in Fig 3. After the random auditory cue was given, a start sound was presented 3 s later, and then the subjects attended the corresponding visual stimuli for 5 s. The auditory cue was given in random order to prevent potentially biased results for the stimulation frequency, the start sound gave the subjects time to prepare to focus on the visual stimuli. The auditory cues were guided by voice
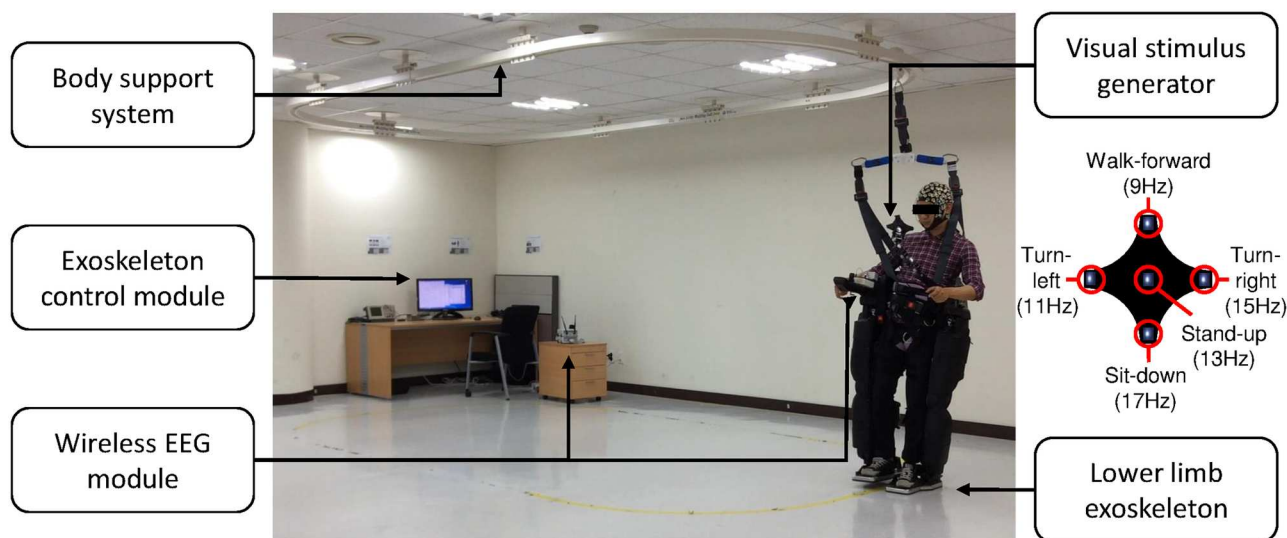

Fig 1. Experimental environment. Subject wearing the lower-limb exoskeleton and focusing on an LED from the visual stimulus generator. The EEG is transferred by a wireless interface to the PC. A body support system (a rail beneath the ceiling) was connected to the subject for safety purposes. The exoskeleton was controlled by an external operator using a keyboard controller.
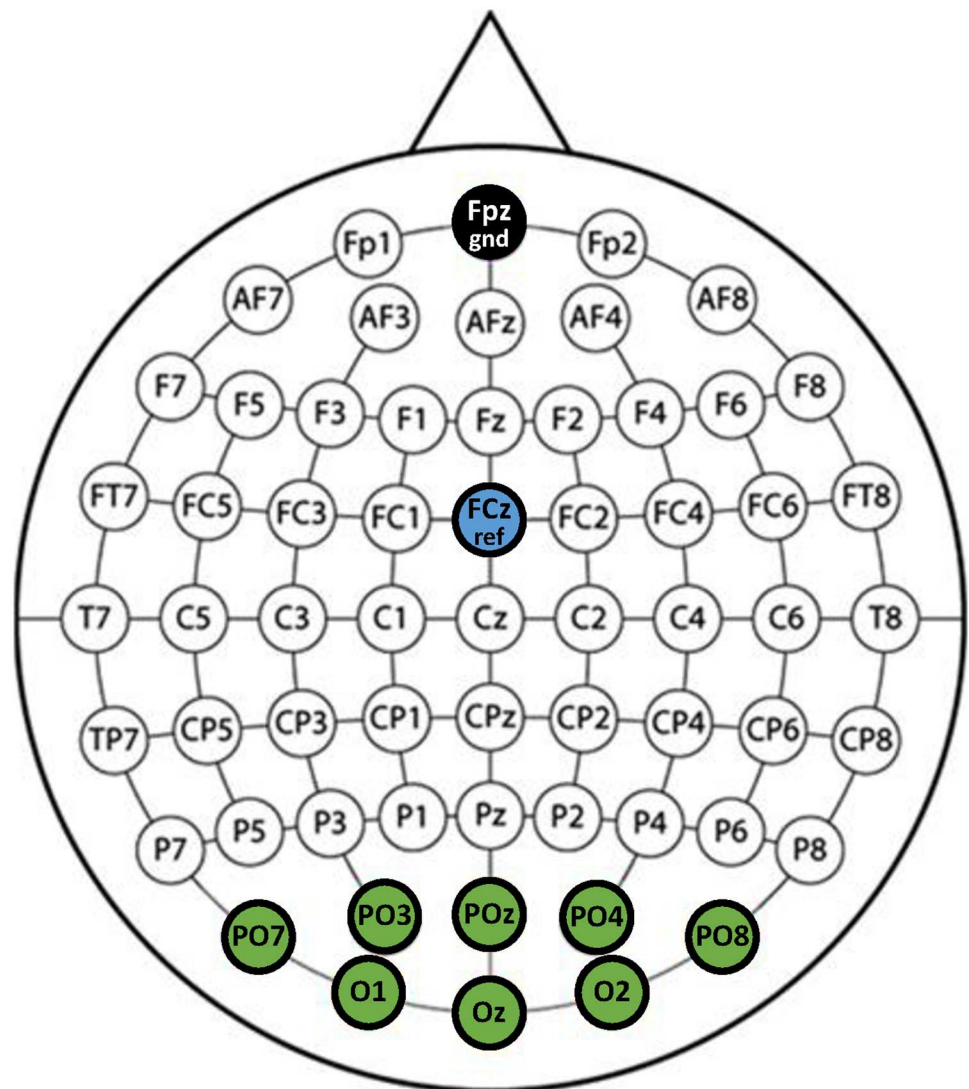
Fig 2. EEG channel layout. Channel layout using 8 channels (PO7, PO3, PO, PO4, PO8, O1, Oz, and O2) for SSVEP acquisition with a reference (FCz) and ground (Fpz).

recordings to indicate commands such as "walk forward", "turn left", "turn right", "sit", and "stand", and were approximately 1 s in length. Note that during the experimental tasks, all LEDs were blinking simultaneously at different frequencies.

- Task 1 (Static SSVEP): The subjects were asked to focus their attention on the visual stimulus in a standing position while wearing the exoskeleton. Corresponding visual stimuli were given by auditory cue and 50 auditory cues were presented in total (10 times in each class).

- Task 2 (Ambulatory SSVEP): The subjects were asked to focus on visual stimuli while engaged in continuous walking using the exoskeleton. The exoskeleton was operated by a wireless controller, per the decoded intention of the subject. a total of 250 auditory cues were presented (50 in each class).
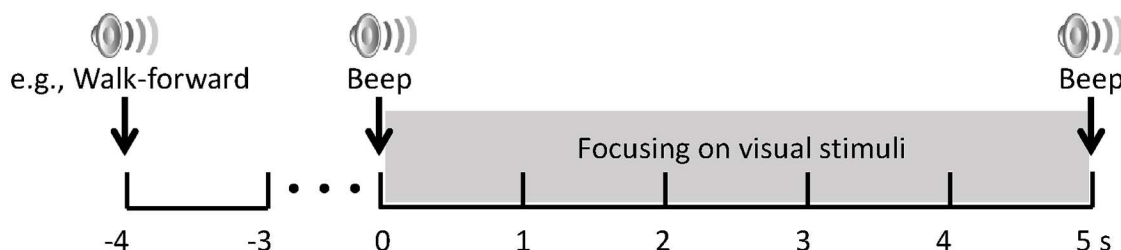
Fig 3. Experimental procedure for Task 1 and 2. After a random auditory cue, a start sound follows 3 s later; then, the subjects attended the corresponding LED for 5 s. All LEDs blinked at differing frequencies during the tasks.

Neural network architectures

We now investigate three neural network architectures for SSVEP decoding, CNN-1 and CNN-2, which use convolutional kernels, and NN, standard feedforward neural network without convolution layers.

We show that CNN-1 has the best classification rate; in CNN-2, we included a fully connected layer with 3 units for visualizing feature representations as a function of the learning progress.

Input data. The acquired EEG data were preprocessed for CNN learning by band-pass filtering from 4–40 Hz. Then, the filtered data were segmented using a 2 s sliding window (2,000 time samples × 8 channels). The segmented data were transformed using a fast Fourier transform (FFT). Then, we used 120 samples from each channel, corresponding to 5–35 Hz. Finally, data were normalized to the range from 0 to 1. Therefore, the input data dimension for CNN learning was 120 frequency samples (N_{fs}) by 8 channels (N_{ch}). The number of input data for training depends on the experimental task and is therefore described in the Evaluation section.

Network architecture overview. The **CNN-1** network has three layers, each composed of one or several maps that contain frequency information for the different channels (similar to [55]). The input layer is defined as $I_{p,\,j}$ with $1 \leq p \leq N_{fs}$ and $1 \leq j \leq N_{ch}$; here, $N_{fs} = 120$ is the number of frequency samples and $N_{ch} = 8$ is the number of channels. The first and second hidden layers are composed of N_{ch} maps. Each map in C_1 has size N_{fs}; each map in C_2 is composed of 110 units. The output layer has 5 units, which represent the five classes of the SSVEP signals. This layer is fully connected to C_2 as in Fig 4.

The **CNN-2** network is composed of four layers. The input layer is defined as $I_{p,\,j}$ with $1 \leq p \leq N_{fs}$ and $1 \leq j \leq N_{ch}$. The first and second hidden layers are composed of N_{ch} maps. Each map in C_1 has size N_{fs}. Each map of C_2 has 110 units. To this point, CNN-2 is equivalent to CNN-1. The difference comes in the third hidden layer F_3, which is fully connected and consist of 3 units. The each unit is fully connected to C_2. The output layer has 5 units that represent the five classes of SSVEP. This layer is fully connected to F_3. The 3 units in F_3 are used to visualize the properties of the representation that CNN-2 has learned, as depicted in Fig 5.

The standard **NN** is composed of three layers. For the input layer, we concatenated the 120 by 8 input into a 960-unit vector. The first hidden layer is composed of 500 units, the second has 100 units, and the output layer has 5 units to represent the five classes. All layers are fully connected, as in Fig 6.

Learning. A unit in the network is defined by $x_k^l(p)$, where l is the layer, k is the map, and p is the position of the unit in the map,

$$x_k^l(p) = f(\sigma_k^l(p)), \qquad (1)$$

where f is the classical sigmoid function used for the layers:

$$f(\sigma) = \frac{1}{1 + exp^{-\sigma}}.\tag{2}$$

$\sigma_k^l(p)$ represents the scalar product of a set of input units and the weight connections between these units and the unit number of p in map k in layer l. For C_1 and C_2, which are convolutional layers, each unit of the map shares the same set of weights. The units of these layers are connected to a subset of units fed by the convolutional kernel from the previous layer. Instead of learning one set of weights for each unit, where the weights depend on unit position, the weights are learned independently to their corresponding output unit. L_3 is the output layer in CNN-1 and L_4 is the output layer in CNN-2.

- CNN-1

    — For C_1:

$$\sigma_k^1(p) = w(1,k,0) + \sum_{j=1}^{j \leq N_{ch}} I_{p,j} w(1,k,j),\tag{3}$$

where $w(1, k, 0)$ is a bias and $w(1, k, j)$ is a set of weights with $1 \leq j \leq N_{ch}$. In this layer, there are N_{ch} weights for each map. The convolution kernel has a size of $1 \times N_{ch}$.

    — For C_2:

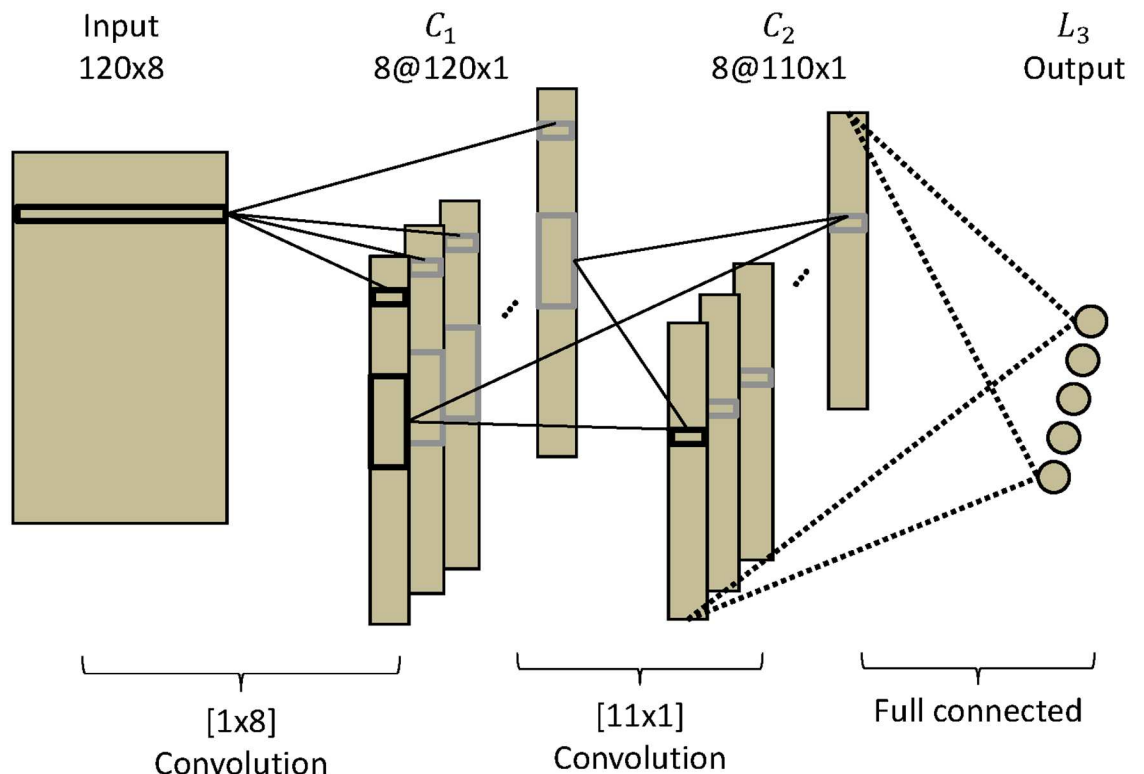$$\sigma_k^2(p) = w(2,k,0) + \sum_{i=1}^{i \leq N_{ch}}\sum_{j=1}^{j \leq 11} x_i^1(p+j-1)w(2,i,j),\tag{4}$$

Fig 4. CNN-1 architecture. CNN-1 is composed of three layers, two convolutional layers and an output layer.
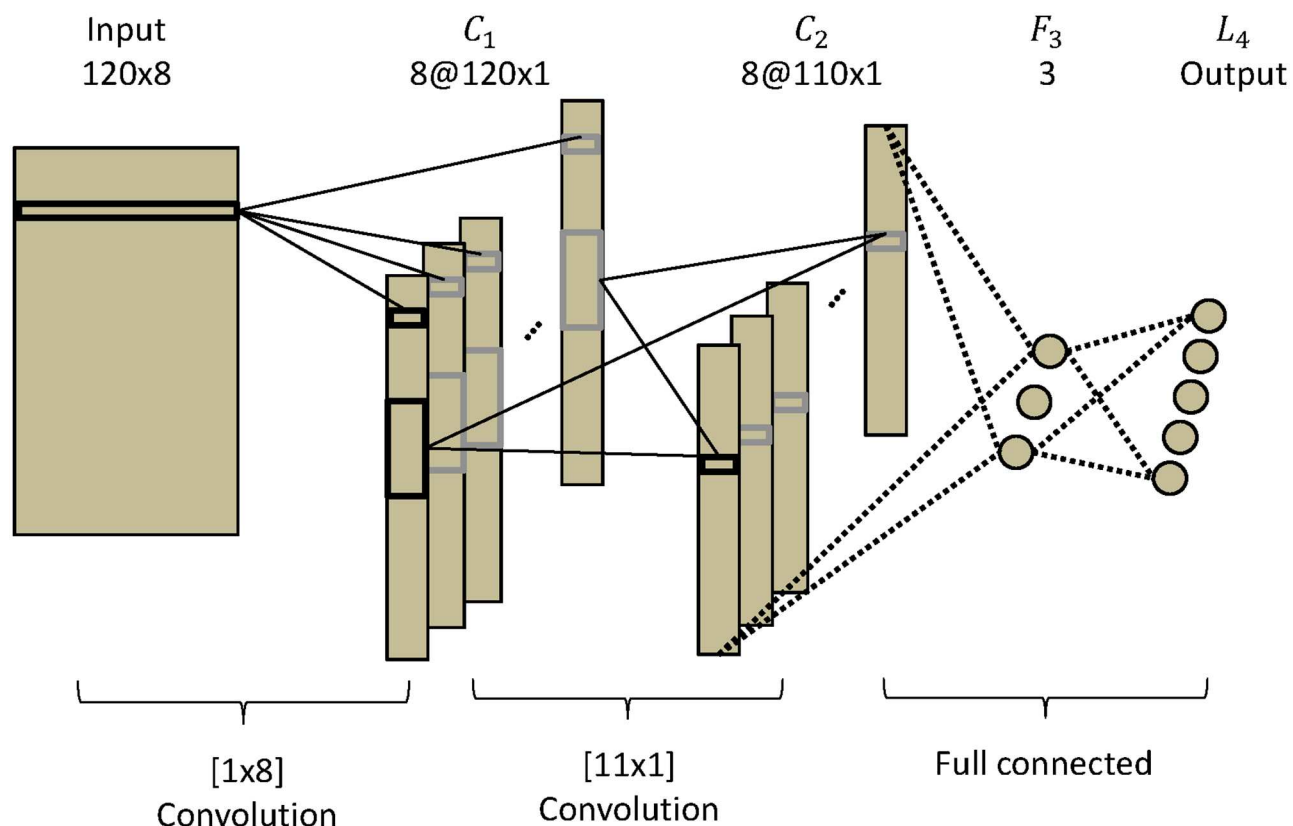
Fig 5. CNN-2 architecture. CNN-2 is composed of four layers: two convolutional layers, one fully connected layer, and an output layer.

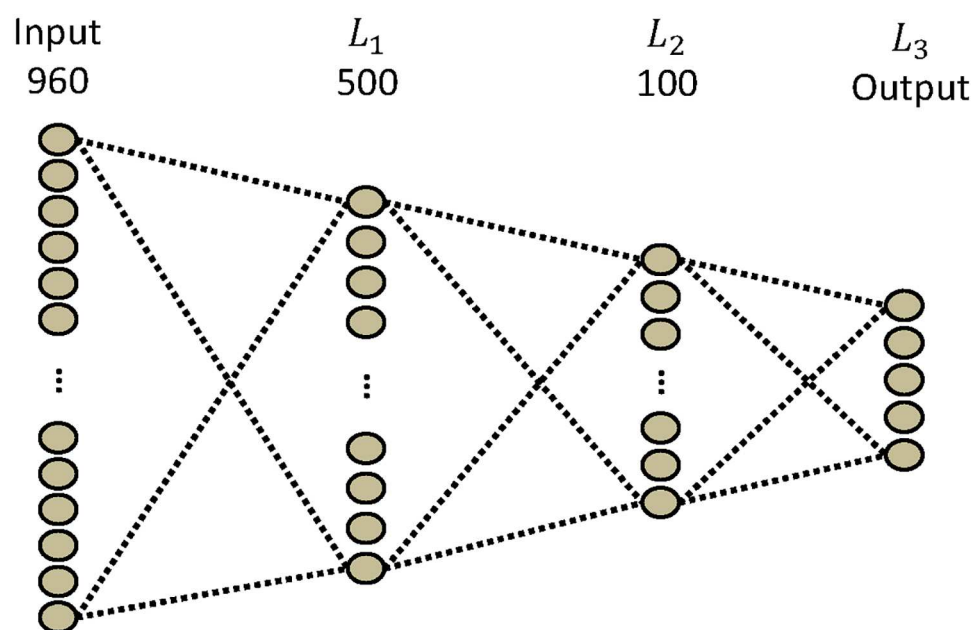doi:10.1371/journal.pone.0172578.g005

Fig 6. NN architecture. The NN is composed of three layers: two fully connected layers and an output layer.

doi:10.1371/journal.pone.0172578.g006

where $w(2, k, 0)$ is a bias. This layer transforms the signal of 120 units into 110 new values in C_2, reducing the size of the signal to analyze while applying an identical linear transformation to the 110 units of each map. This layer translates spectral filters. The convolution kernel has a size of 11×1.

— For L_3:

$$\sigma^3(p) = w(3, 0, p) + \sum_{k=1}^{k \leq 8} \sum_{l=1}^{l \leq 110} x_k^2(l) w(3, k, l), \tag{5}$$

where $w(3, 0, p)$ is a bias. Each unit of L_3 is connected to each unit of C_2.

- CNN-2

  — C_1 and C_2 are the same as in CNN-1.

  — For F_3:

$$\sigma^3(p) = w(3, 0, p) + \sum_{k=1}^{k \leq 8} \sum_{l=1}^{l \leq 110} x_k^2(l) w(3, k, l), \tag{6}$$

where $w(3, 0, p)$ is a bias. Each unit of F_3 is connected to each unit of C_2

  — For L_4:

$$\sigma^4(p) = w(4, 0, p) + \sum_{l=1}^{l \leq 3} x^3(l) w(4, l), \tag{7}$$

where $w(4, 0, p)$ is a bias. Each unit of L_4 is connected to each unit of F_3

The gradient descent learning algorithm uses standard error backpropagation to correct the network weights [56–58]. The learning rate was 0.1 and weights were initialized with a normal distribution on the interval [-sqrt(6/(N_{in}+N_{out})), sqrt(6/(N_{in}+N_{out}))], where N_{in} is the number of input weights and N_{out} is the number of output weights following [58]. The number of learning iterations was 50, but training stopped once the decrease of in error rate was smaller than 0.5% after 10 iterations.

Evaluation

In this section, we validate the three neural networks (CNN-1, CNN-2, and NN), and compare them to previously used methods: CCA [37, 48], MSI [43] and CCA combined with k-nearest neighbors (CCA-KNN) [14].

For each classifier, we compute the 10-fold cross-validation error, splitting the data chronologically (a common method in EEG classification) to preserve the data's non-stationarity and avoid overfitting [27, 59]. For the test data, both datasets (50 trials of 5 s for static SSVEPs and 250 trials for ambulatory SSVEPs, randomly permuted) were segmented using a 2 s sliding window with a 10 ms shift size, segmenting a 5 s trial into three hundred 2 s trials. As a result, there were 1,500 static and 7,500 ambulatory SSVEP test data points in each fold. Deep neural networks generally show higher performance for larger amounts of data [53]. Hence, we tested the classifiers with different training data sizes; in particular, different segmentations of the data were considered. Using a 2 s sliding window with different shift sizes (60, 30, 20, 15, 12, and 10 ms), we obtained a trial segmentation into 50, 100, 150, 200, 250, and 300 data samples. Thus, there were 2,250, 4,500, 6,750, 9,000, 11,250, and 13,500 training data for the static SSVEPs, and 11,250, 22,500, 33,750, 45,000, 56,250, and 67,500 for the ambulatory SSVEPs.

Note that although we used a small size shift, there was no overlap between training and test data in order to prevent overfitting. The CCA method does not require a training phase. Thus, we only show its results on test data.

We now briefly describe the CCA, CCA-KNN, and MSI methods. CCA is a multivariate statistical method [60, 61] that finds a pair of linear combinations such that the correlation between two canonical variables X and Y is maximized. As $X(t)$, we chose 2 s EEG windows; as $Y_i(t)$, we use the five reference frequencies ($f_1 = 9, f_2 = 11, \ldots, f_5 = 17$) from the five visual stimuli [14]

$$Y_i(t) = (sin(2\pi f_i t), cos(2\pi f_i t), sin(2\pi(2f_i)t), cos(2\pi(2f_i)t))', \quad t = \frac{1}{S}, \frac{2}{S}, \cdots, \frac{T}{S}, \quad (8)$$

where T is the number of sampling points and S denotes the sampling rate. CCA finds weight vectors, W_x and W_y, that maximize the correlation between the canonical variants $x = X'W_x$ and $y = Y'W_y$, by solving

$$\max_{W_x W_y} \rho(x, y) = \frac{E[x'y]}{\sqrt{E[x'x]E[y'y]}} = \frac{E[W_x'XY'W_y]}{\sqrt{E[W_x'XX'W_x]E[W_Y'YY'W_y]}}. \quad (9)$$

The maximum ρ with respect to W_x and W_y is the maximum canonical correlation. The canonical correlation ρ_{f_i}, where i = 1, . . ., 5, is used for detecting the frequency of the LED that a subject is attending by

$$O_i = \max_i(\rho_{f_i}), \quad i = 1, \ldots, 5, \quad (10)$$

where O_i are the output classes corresponding to the five visual stimuli.

For CCA-KNN [14], the set of canonical correlations ($\boldsymbol{\rho} = (\rho_{f_1}, \ldots, \rho_{f_5})$) is used as a feature vector for subsequent KNN classification, each with a class label. In the training step, the algorithm consists only of storing the feature vectors and class labels of the training samples. In classification, an unlabeled vector is classified by assigning the label to the most frequent of the k nearest training samples, where the Euclidean distance is used as a distance metric.

For MSI [43, 44], the S-estimator, based on the entropy of the normalized eigenvalues of the correlation matrix of multivariate signals, was used as the index. Thus, MSI creates a reference signal from the stimulus frequencies used in an SSVEP-based BCI system similarly to CCA.

Results and discussion

EEG signals are highly variable across subjects and experimental environments (see Figs 5 and 6 and Tables 1 and 2 in [14]). The SSVEP signals acquired from the static exoskeleton show more pronounced frequency information than in the ambulatory environment. In the static SSVEP, we can observe the increased frequency components that are visible at the stimulus frequency. In the ambulatory SSVEP, however, because of the higher artifactual content, this effect becomes less clearly visible (see S1 Fig for selected input and average data under both conditions).

Static SSVEP

In Table 1, we show the 10-fold cross-validation results for 13,500 training data validated on 1,500 test data points for all subjects. CNN-1 showed the best classification accuracy of all subjects in each classifier. For low-performing subjects, with a CCA accuracy under 80% in the

Table 1. 10-fold cross validation results of individual subjects with the maximum quantity of training data (i.e., 13,500) using static SSVEP.

	S1	S2	S3	S4	S5	S6	S7	Low	All
CCA	91.27 (-8.59)	93.78 (-6.84)	76.14 (-22.72)	92.86 (-6.3)	67.85 (-31.62)	78.22 (-20.39)	82.56 (-16.77)	79.53±9.17 (-19.55)	83.24±9.84 (-16.04)
MSI	91.54 (-8.32)	95.71 (-3.99)	78.74 (-20.12)	94.67 (-4.49)	70.61 (-28.85)	79.34 (-19.27)	83.43 (-15.91)	81.86±8.77 (-17.73)	84.86±9.41 (-14.42)
CCA-KNN	100±0 (+0.14)	99.80±0.27 (+0.1)	93.91±2.11 (-4.95)	98.97±1.11 (-0.19)	99.13±0.95 (-0.34)	93.26±3.70 (-5.35)	98.86±0.91 (-0.47)	96.83±2.97 (-2.25)	97.70±2.85 (-1.58)
NN	98.95±1.51 (-1.01)	98.85±1.9 (-0.85)	97.53±2.95 (-1.33)	98.87±1.28 (-0.29)	99.73±0.70 (+0.26)	97.11±3.42 (-1.5)	98.01±2.68 (-1.32)	98.25±1.05 (-0.64)	98.44±0.92 (-0.84)
CNN-2	98.46±1.97 (-1.4)	98.17±2.16 (-1.53)	97.90±2.24 (-0.96)	97.43±2.15 (-1.73)	99.51±0.70 (+0.04)	96.54±3.18 (-2.07)	97.50±2.75 (-1.83)	97.63±1.54 (-1.45)	97.83±1.31 (-1.45)
CNN-1	99.86±0.27	99.70±0.27	98.86±1.33	99.16±1.77	99.47±0.76	98.61±2.30	99.33±0.72	99.08±0.35	99.28±0.45

10-fold cross validation results of static SSVEP classification for 13,500 training data points with 1,500 test data for all subjects. Low indicates subjects who have a low CCA accuracy (under 80% in the ambulatory SSVEP, i.e., subjects S3–7). Parentheses indicate accuracy the difference compared with CNN-1.

ambulatory SSVEP (i.e., subjects S3–7, see Table 3), the neural network results stayed robust. Clearly, the CCA method exhibits significantly lower performance.

With fewer training data (see Table 2 and Fig 7 (top) for the 10-fold cross-validation results of the static task), we observe a decaying performance for the neural networks, which is to be expected. Note that the CCA and MSI methods stay essentially constant as a function of data, since no training phase is required because the canonical correlations and synchronization index with reference signals are simply computed in order to find the maximum value. The CCA-KNN classifier was trained for $k = 1, 3, 5,$ and 7, respectively, and k was selected on the training set to achieve the best accuracy. Fig 7 (top) presents the average accuracy of each classifier as a function of the number of training data for all subjects (a) and low-performing subjects (b). Statistical analysis of these results shows a significant improvement with larger training data sizes for the neural network classifiers. We provide more information on the difference between CNN-1 and the other methods for all subjects and low-performing subjects in S2(a) and S2(b) Fig, respectively. CNN-1 outperforms other classifiers; however, CCA-KNN shows better classification results for 4,500 training data samples or fewer, as we can see from the positive values in brackets.

Fig 8(a) shows the decoding variability of the individual subjects at the minimum (dash) and maximum (solid line) number of data samples; here, a diamond indicates a 5% or more

Table 2. 10-fold cross-validation of static SSVEP classification by the quantity of training data.

	2,250		4,500		6,750		9,000		11,250		13,500	
	Low	All	Low	All	Low	All	Low	All	Low	All	Low	All
CCA-KNN	96.85 (+1.76)	97.71 (+1.51)	97.91 (+0.38)	98.47 (+0.48)	96.83 (-0.29)	97.60 (-0.07)	96.79 (-1.23)	97.57 (-0.78)	96.83 (-1.61)	97.59 (-1.13)	96.83 (-2.25)	97.70 (-1.58)
NN	91.82 (-3.27)	92.85 (-3.35)	96.28 (-1.25)	96.55 (-1.44)	97.00 (-0.12)	97.29 (-0.38)	97.61 (-0.41)	97.81 (-0.54)	98.04 (-0.4)	98.21 (-0.51)	98.25 (-0.64)	98.44 (-0.84)
CNN-2	90.49 (-4.6)	91.16 (-5.04)	95.74 (-1.79)	96.11 (-1.88)	93.98 (-3.14)	94.89 (-2.78)	96.45 (-1.57)	96.73 (-1.62)	96.78 (-1.66)	97.27 (-1.45)	97.63 (-1.45)	97.83 (-1.45)
CNN-1	95.09	96.20	97.53	97.99	97.12	97.67	98.02	98.35	98.44	98.72	99.08	99.28

10-fold cross-validation of static SSVEP classification, changing the amount of training data with 1,500 test data. Low indicates subjects who have a low CCA accuracy (under 80% in the ambulatory SSVEP, i.e., subjects S3–7). Parentheses indicate the differences in accuracy when compared with CNN-1.

Table 3. 10-fold cross-validation results of individual subjects at the maximum training data (i.e., 67,500) for ambulatory SSVEP classification.

	S1	S2	S3	S4	S5	S6	S7	Low	All
CCA	92.43 (-5.24)	90.40 (-8.41)	73.10 (-18.3)	65.46 (-22.05)	43.99 (-47.99)	69.90 (-23.48)	75.05 (-21.85)	65.50±12.56 (-26.85)	72.90±16.29 (-21.13)
MSI	94.02 (-3.65)	92.35 (-6.46)	78.31 (-13.09)	69.22 (-18.29)	43.67 (-48.30)	71.00 (-22.95)	77.80 (-19.11)	68.00±14.18 (-24.35)	75.19±16.89 (-18.84)
CCA-KNN	97.69±0.82 (+0.02)	97.53±0.78 (-1.28)	82.77±3.05 (-8.63)	77.27±3.82 (-10.24)	68.25±3.76 (-23.73)	85.43±2.57 (-7.95)	83.58±2.05 (-13.32)	79.46±6.96 (-12.87)	84.65±10.52 (-9.38)
NN	94.77±3.48 (-2.9)	95.11±3.58 (-3.7)	91.03±6.82 (-0.37)	90.20±7.82 (+2.69)	91.35±6.51 (-0.63)	92.07±6.54 (-1.31)	95.03±3.54 (-1.87)	91.94±1.86 (-0.41)	92.80±2.11 (-1.23)
CNN-2	95.76±4.47 (-1.91)	96.99±2.51 (-1.82)	83.36±5.16 (-8.04)	80.09±4.4 (-7.42)	84.96±3.71 (-7.02)	87.32±4.92 (-6.06)	90.07±2.42 (-6.83)	85.16±3.80 (-7.19)	88.36±6.30 (-5.67)
CNN-1	97.67±1.50	98.81±0.74	91.40±3.21	87.51±4.8	91.98±3.23	93.38±2.65	96.90±1.92	92.35±3.46	94.03±4.04

10-fold cross-validation results of ambulatory SSVEP classification for 67,500 training data and 7,500 test data. Low indicates subjects who have a low CCA accuracy (under 80%, i.e., subjects S3–7). In parentheses, the accuracy differences compared with CNN-1 are indicated.

doi:10.1371/journal.pone.0172578.t003

increase in classification rates. Clearly, all subjects achieved increased accuracies with neural network models. Specifically, low-performing subjects (S3–7) show a higher increase than other subjects (S1 and S2). Subjects S2 and S3 only increase in the CCA-KNN method.

Ambulatory SSVEP

Table 3 considers the 10-fold cross-validation results for the ambulatory SSVEP setup when the number of training data is 67,500, with 7,500 test data for all subjects. CNN-1 showed the best classification accuracy of all subjects in each classifier. Even the low-performing subjects showed the highest accuracy with CNN-1. The competing classifier models showed a more pronounced performance deterioration owing to the higher artifact presence in the ambulatory environment (see Figs 5 and 6 in [14]).

Table 4 and Fig 7 (bottom) presents the 10-fold cross-validation results for the ambulatory SSVEP classification as a function of a changing number of training data with 7,500 test data. In particular, Fig 7 (bottom) shows the averaged accuracy of each classifier with increasing training data for all subjects (c) and low-performance subjects (d). One asterisk indicates the 5% significance level (compared to 67,500 training data samples), whereas two asterisks denote the 1% significance level. As expected, analysis confirms the performance gains of the neural networks as training data increases, even if the data contain large artifacts. Fig 8(b) confirms the findings of the static setting for the more artifact-prone ambulatory setting.

Compared with the static SSVEP in CNN-1, larger training data samples are required for the ambulatory SSVEP to accomplish high accuracy (classification performance of more than 90% at 56,250 training data samples). For the static SSVEP setup, 96.20% accuracy could be achieved using only 2,250 training data and 99.28% accuracy was achieved using 13,500 samples. In contrast, 81.40% accuracy and 94.03% accuracy were achieved in the ambulatory condition when 11,250 and 67,500 training data were used, respectively.

Fig 9 shows the learning curves of subjects S2 (black) and S4 (red) in static (solid) and ambulatory (dash line) SSVEP environments in CNN-1. The learning iteration of subject S2 stops at the 13th and 12th epochs, whereas the iteration of subject S4 stops at the 19th and 12th epochs in the datasets (subject S2 records the best performance with CNN-1 and subject S4 has the lowest performance in the ambulatory SSVEP.)

The appearance of the kernels differs for each individual because the network training is subject-dependent. Unfortunately, there is no obvious and simple interpretation linked to
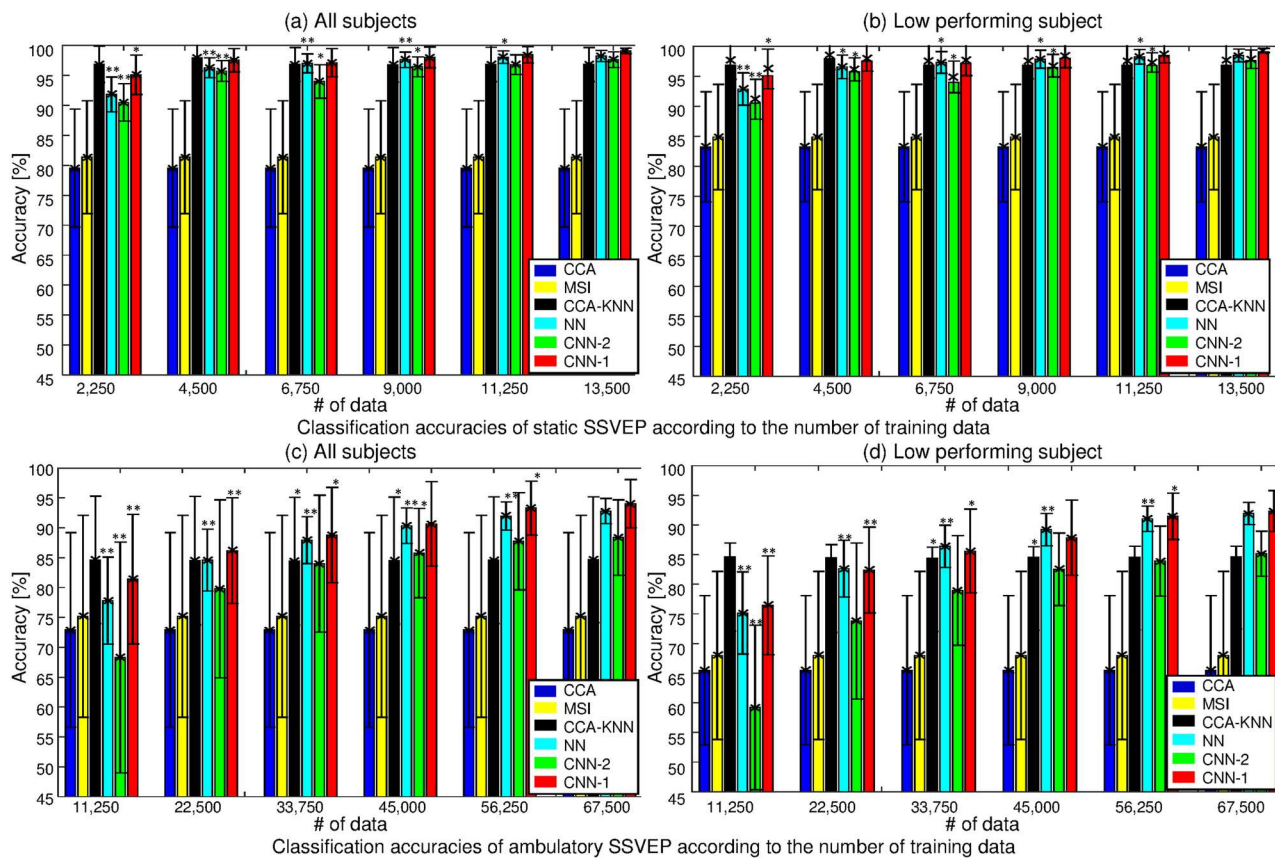
Fig 7. 10-fold cross-validation performance comparison using static (top) and ambulatory SSVEP (bottom) as the number of training data increases. (a) Average of all subjects in static SSVEP. (b) Low performing subjects. 1,500 test data were used for each fold. One asterisk indicates indicates the 5% significance level between corresponding to the number of training data samples and 13,500 training data samples. Two asterisks are the 1% significance level. (c) Average of all subjects in ambulatory SSVEP. (d) Low performing subjects. 7,500 test data were used for each fold. One asterisk indicates indicates the 5% significance level between the corresponding number of training data samples and 67,500 training data samples. Two asterisks are the 1% significance level.

doi:10.1371/journal.pone.0172578.g007
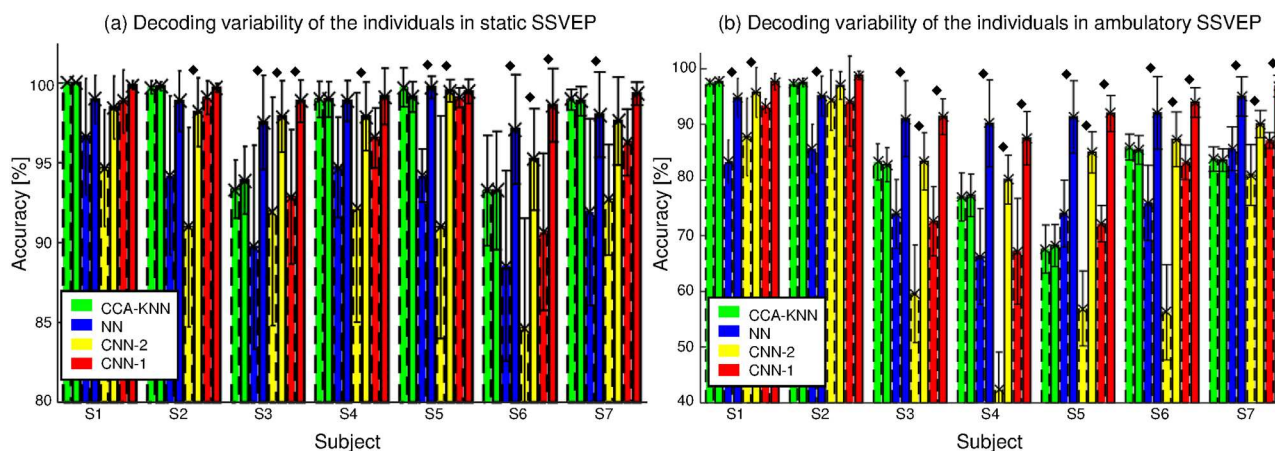
Fig 8. Decoding variability across individuals. Decoding variability of individuals at the minimum and maximum number of data samples (dash and solid line, respectively) in static (a) and ambulatory (b) SSVEP environments.

doi:10.1371/journal.pone.0172578.g008

Table 4. 10-fold cross-validation of ambulatory SSVEP classification by the quantity of training data.

	11,250		22,500		33,750		45,000		56,250		67,500	
	Low	All	Low	All	Low	All	Low	All	Low	All	Low	All
CCA-KNN	79.51 (+3.04)	84.64 (+3.24)	79.28 (-3.11)	84.49 (-1.69)	79.16 (-6.43)	84.41 (-4.38)	79.29 (-8.56)	84.52 (-6.14)	79.36 (-12.09)	84.58 (-8.72)	79.46 (-12.87)	84.65 (-9.38)
NN	75.15 (-0.88)	77.81 (-3.59)	82.63 (+0.24)	84.61 (-1.57)	86.39 (-0.8)	87.93 (-0.86)	89.21 (+1.36)	90.35 (-0.31)	91.04 (-0.41)	91.96 (-1.34)	91.94 (-0.41)	92.80 (-1.23)
CNN-2	59.21 (-17.26)	68.31 (-13.09)	73.77 (-8.62)	79.78 (-3.4)	78.95 (-6.64)	83.99 (-4.8)	82.52 (-5.33)	85.77 (-4.89)	83.91 (-7.54)	87.75 (-5.55)	85.16 (-7.19)	88.36 (-5.67)
CNN-1	76.47	81.40	82.39	86.18	85.59	88.79	87.85	90.66	91.45	93.30	92.35	94.03

10-fold cross-validation of the ambulatory SSVEP classification when changing the number of training data with 7,500 test data. Low indicates subjects who have a low CCA accuracy (under 80%, i.e., subjects S3–7). In parentheses, the accuracy differences of the methods compared with CNN-1 are indicated.

doi:10.1371/journal.pone.0172578.t004

physiology or the experimental task (see S3 Fig which describes the convolutional kernels learned from CNN-2 using ambulatory SSVEP data for subject S2 (top), and subject S3 (bottom)).

Fig 10 shows the decoding trends of CNN-1 compared with CCA-KNN for individuals as a function of the number of training data in the static (a) and ambulatory (b) SSVEP setups. The darker circles indicate more training data. As more training data were given, we observed that the performance of the CNN-1 increased consistently and was more pronounced under the ambulatory condition (more black circles are located on the right side). However, individual CCA-KNN decoding accuracies stay relatively stable, meaning that the accuracy of the CCA-KNN is almost independent of the amount of training data in our experimental conditions.
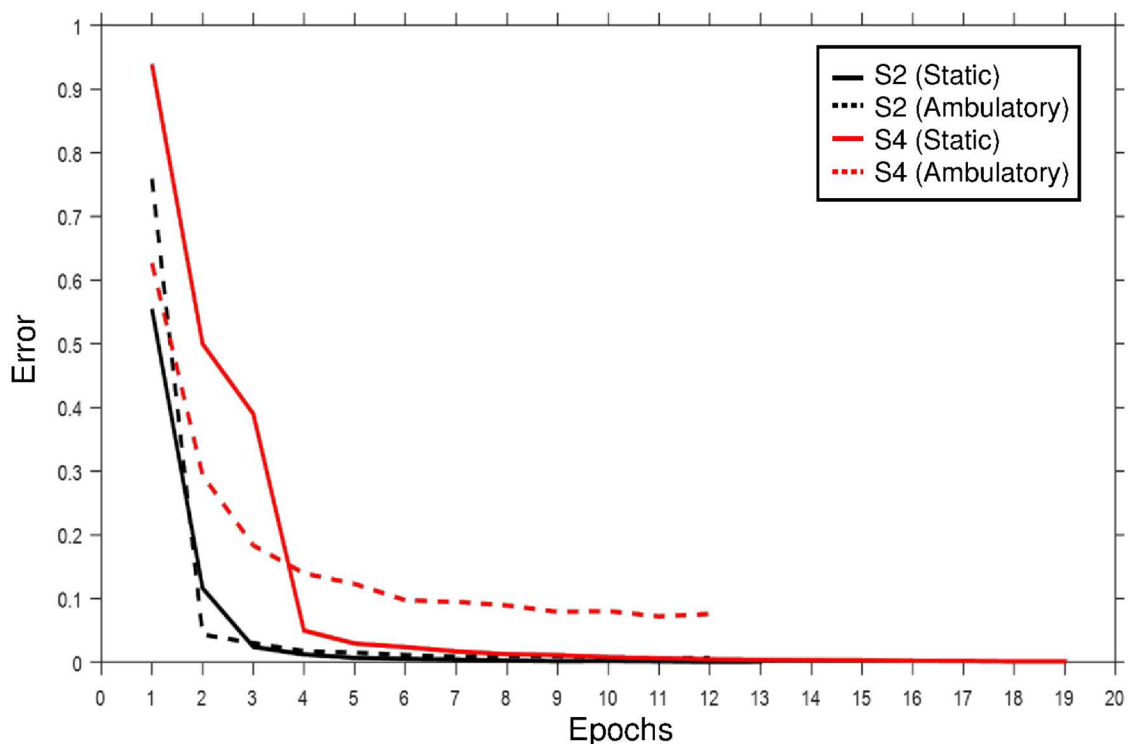

Fig 9. Learning curve for subjects S2 and S4. Learning curve of subjects S2 (black) and S4 (red) using static (solid) and ambulatory (dash line) SSVEP environments in CNN-1.
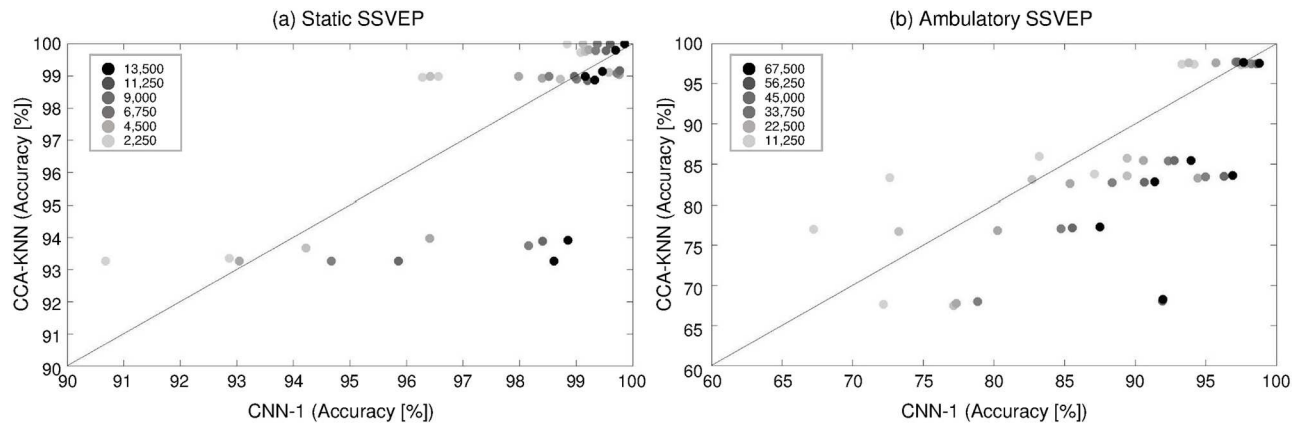
doi:10.1371/journal.pone.0172578.g009

Fig 10. Decoding trends of CNN-1 compared with CCA-KNN for the individual subjects. Decoding trends of CNN1 compared with CCA-KNN by the number of training data in static SSVEP (a) and ambulatory (b) SSVEP environments. The circles indicate a higher quantity of training data with darker color. The more training data were given, the better CNN-1 performed (more black circles are located on the right side).

doi:10.1371/journal.pone.0172578.g010

Feature representation

Analyzing and understanding classification decisions in neural networks is valuable in many applications, as it allows the user to verify the system's reasoning and provides additional information [54, 62]. Although deep learning methods are very successfully solving various pattern recognition problems, in most cases, they act as a black box, not providing any information about why a particular decision was made. Hence, we present the feature representation from the CNN-2 architecture. The averaged features of each layer using static and ambulatory SSVEPs are shown in Figs 11 and 12, respectively. In both cases, the networks focus on the stimulus frequency components. For learning in layer C_1, we used a $1 \times N_{ch}$
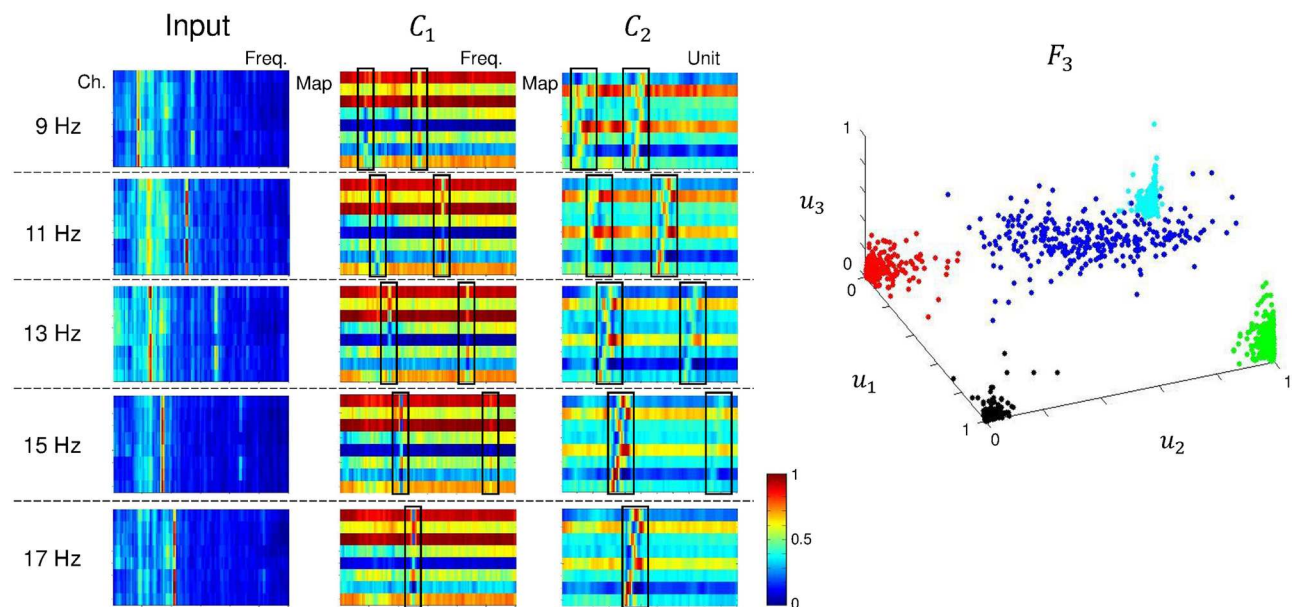

Fig 11. Feature representation of CNN-2 using static SSVEPs for subject S1. Representation of the average features of each layer in CNN-2 using static SSVEP data. In layer F_3, blue is 9 Hz; red, 11 Hz; green, 13 Hz; black, 15 Hz; and cyan, 17 Hz.
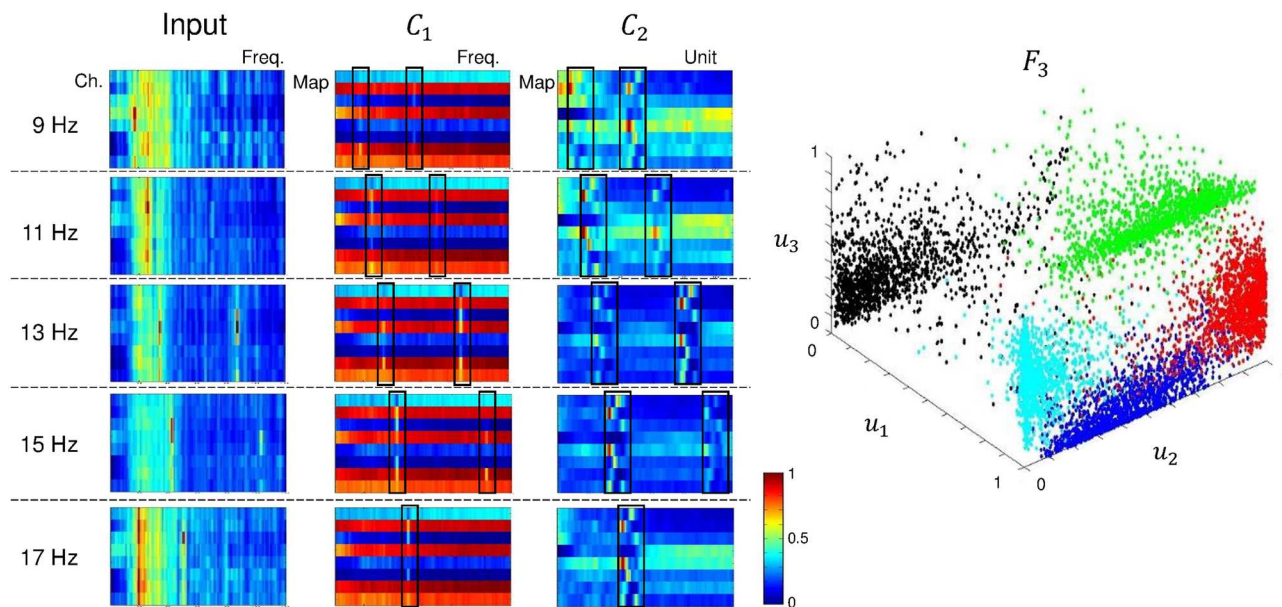
doi:10.1371/journal.pone.0172578.g011

Fig 12. Feature representation of CNN-2 using ambulatory SSVEP for subject S1. Representation of the average features of each layer in CNN-2 using ambulatory SSVEP data. In layer F_3, blue is 9 Hz; red, 11 Hz; green, 13 Hz; black, 15 Hz; and cyan, 17 Hz.

doi:10.1371/journal.pone.0172578.g012

convolutional kernel, which can give channel-wise (spatial) weight. The C_2 layer used an 11×1 convolutional kernel to detect frequency (spectral) information. The frequency components that were most discriminated by the convolutional layers were highlighted using black-lined boxes. With the exception of the 17 Hz class, the corresponding stimulus frequencies were enforced through iterative training. We conjecture that the absence of second harmonics (34 Hz) for the 17 Hz SSVEPs results from low magnitude when compared with lower frequencies or outside the boundary of the ranges in the C_2 layer. In the second convolutional layer, the patterns were spread out (and slightly smoothed) when compared to the first convolutional layer. The F_3 layer is composed of three units that we plotted with each unit as an axis direction. The 3D plot shows that all classes are distinguished nicely. Therefore, we can conclude that the CNN architecture is able to appropriately extract the meaningful frequency information of SSVEP signals. To compare the feature distributions with CCA-KNN, we show a scatter plot using CCA-KNN in S4 Fig. The features were extracted with CCA and classified using KNN when $k = 3$ for subject S6 (85%). Test data were plotted on ρ_{f_1}, ρ_{f_2} and ρ_{f_3} axes. Blue, red, green, black, and cyan circles indicate 9 Hz, 11 Hz, 13 Hz, 15 Hz, and 17 Hz, respectively. Note that the feature dimension is actually 5 (the number of classes), therefore we only used the ρ_{f_1}, ρ_{f_2} and ρ_{f_3} projection to visualize feature distributions in the plot. However, the classes are clearly not as well spread apart when compared with CNN-2.

Conclusion

BMI systems have shown great promise, though significant effort is still required to bring neuroprosthetic devices from the laboratory into the real world. In particular, further advancement in the robustness of brain signal processing techniques is needed [63, 64]. In this context, constructing reliable BMI-based exoskeletons is a difficult challenge owing to the various complex artifacts spoiling the EEG signal. These artifacts may be induced differently depending on subject population and may in particular be caused by suboptimal EEG measurements or broadband distortions due to movement of the exoskeleton. For example, while walking in the

exoskeleton a subject's head may move, which can give rise to swinging movements in the line between the electrodes and EEG amplifiers, leading to disconnections or high impedance measurements. Furthermore, significant challenges still exist in the development of a lower-limb exoskeleton that can integrate with the user's neuromusculoskeletal system. Although these limitations exist, a brain-controlled exoskeleton may eventually be helpful for end-user groups.

The current study made a step forward toward more robust SSVEP-BMI classification. Despite the challenges imposed on signal processing by a lower-limb exoskeleton in an ambulatory setting, our proposed CNN exhibited promising and highly robust decoding performance for SSVEP signals. The neural network model was successfully evaluated offline against standard SSVEP classification methods on SSVEP datasets from static and ambulatory tasks. The three neural networks (CNN-1, CNN-2, and NN) showed increased performance in both environments when sufficient training data were provided. CNN-1 outperformed all other methods; the best accuracies achieved by CNN-1 were 99.28% and 94.03% in static and ambulatory conditions, respectively. Other methods (CCA-KNN, NN, CNN-2) showed high accuracy in the static environment, but only CNN-1 recorded smallest low performance deterioration for the ambulatory SSVEP task. CNN-1's complexity is low because it has a comparatively simple structure (few layers, maps, and units) and the weights in the convolution layers are shared for every unit within one map, effectively reducing the number of free parameters in the network. Our application is far from being data rich ($N \leq 67,500$); therefore, we adopted neither pre-trained model, dropout, nor pooling methods, yet our relatively simple architecture worked efficiently after a brief training period. Overall, the proposed method has advantages for real-time usage and it is highly accurate in the ambulatory conditions. Furthermore, our method can increase in accuracy with more data, if available. Note that we consider subject-dependent classifiers for decoding, which reflects the fact that individuals possess highly different patterns in their brain signals. From the kernel analysis, we therefore found—as expected—that the convolutional kernels were different for each individual. We also demonstrate the feature representations, as implemented using a bottleneck layer in CNN-2. The CNN classifiers could determine the most discriminative frequency information for classification, nicely matching the stimulus frequencies of the respective SSVEP classes.

So far, our study has only successfully tested the performance of CNN classifiers for offline data. Future work will also develop a real-time CNN system that can control a lower-limb exoskeleton based on the proposed method and evaluate its performance with healthy volunteers as well as for end-user groups to investigate their use in gait rehabilitation. We will investigate subject-independent classification using CNNs. A subject-independent CNN-based classifier may be more efficient system because it could reduce long training times.

Supporting information

S1 Fig. Examples of input data. Randomly selected input data and averaged data of (a) static and (b) ambulatory SSVEPs for a representative subject S7. Red boxes indicate the frequency location corresponding to stimulus frequencies.
(PDF)

S2 Fig. Accuracy differences in 10-fold cross-validation performance using static (top) and ambulatory (bottom) SSVEPs as the number of training data increases. (a) Accuracy differences for all subjects in static SSVEP. (b) Accuracy differences for low-performance subjects in static SSVEP. (c) Accuracy differences for all subjects in ambulatory SSVEP. (d) Accuracy differences for low-performance subjects in ambulatory SSVEP.
(PDF)

S3 Fig. Kernel appearance. Kernels of layer C_1 (left) and C_2 (right) in CNN-2 using ambulatory SSVEPs for S2 (top) and S3 (bottom).
(PDF)

S4 Fig. A feature representation of CCA-KNN. Features were extracted from CCA and classified using KNN with $k = 3$ for subject S6. Test data were plotted along the ρ_{f_1}, ρ_{f_2} and ρ_{f_3} axes. Blue, red, green), black, and cyan are 9, 11, 13, 15, and 17 Hz, respectively.
(PDF)

Acknowledgments

Author Contributions

Formal analysis: NSK.

Investigation: NSK.

Methodology: NSK KRM SWL.

Project administration: SWL.

Supervision: SWL.

Validation: KRM.

Visualization: NSK.

Writing – original draft: NSK KRM SWL.

Writing – review & editing: NSK KRM SWL.

References

1. Lebedev MA, Nicolelis MA. Brain-machine interfaces: past, present and future. TRENDS in Neurosciences. 2006 Sep 30; 29(9):536–546. doi: 10.1016/j.tins.2006.07.004 PMID: 16859758

2. Dornhege G, Millán JR, Hinterberger T, McFarland DJ, Müller K-R. Towards brain-computer interfacing. MIT press; 2007.

3. Müller K-R, Blankertz B. Toward noninvasive brain-computer interfaces. IEEE Signal Processing Magazine. 2006 Sep 1; 23(5):125–128.

4. Donchin E, Spencer KM, Wijesinghe R. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. IEEE Transactions on Rehabilitation Engineering. 2000 Jun; 8(2):174–9. doi: 10.1109/86.847808 PMID: 10896179

5. Hwang H-J, Lim J-H, Jung Y-J, Choi H, Lee S-W, Im C-H. Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. Journal of Neuroscience Methods. 2012 Jun 30; 208(1):59–65. doi: 10.1016/j.jneumeth.2012.04.011 PMID: 22580222

6. Lopes AC, Pires G, Nunes U. Assisted navigation for a brain-actuated intelligent wheelchair. Robotics and Autonomous Systems. 2013 Mar 31; 61(3):245–258. doi: 10.1016/j.robot.2012.11.002

7. Carlson T, Millán JR. Brain-controlled wheelchairs: a robotic architecture. IEEE Robotics and Automation Magazine. 2013; 20(1):65–73. doi: 10.1109/MRA.2012.2229936

8. Tangermann MW, Krauledat M, Grzeska K, Sagebaum M, Vidaurre C, Blankertz B, et al. Playing Pinball with non-invasive BCI. In: NIPS 2008 Dec 8;1641–1648.

9. Krepki R, Blankertz B, Curio G, Müller K-R. The Berlin Brain-Computer Interface (BBCI)-towards a new communication channel for online control in gaming applications. Multimedia Tools and Applications. 2007 Apr 1; 33(1):73–90. doi: 10.1007/s11042-006-0094-3

10. Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, et al. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature. 2006 Jul 13; 442(7099):164–171. doi: 10.1038/nature04970 PMID: 16838014

11. Collinger JL, Wodlinger B, Downey JE, Wang W, Tyler-Kabara EC, Weber DJ, et al. High-performance neuroprosthetic control by an individual with tetraplegia. The Lancet. 2013 Feb 22; 381:557–564. doi: 10.1016/S0140-6736(12)61816-9

12. Kim J-H, Bießmann F, Lee S-W. Decoding three-dimensional trajectory of executed and imagined arm movements from electroencephalogram signals. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2015 Sep; 23(5):867–876. doi: 10.1109/TNSRE.2014.2375879 PMID: 25474811

13. King CE, Wang PT, Chui LA, Do AH, Nenadic Z. Operation of a brain-computer interface walking simulator for individuals with spinal cord injury. Journal of NeuroEngineering and Rehabilitation. 2013 Jul 17; 10(1):1–14. doi: 10.1186/1743-0003-10-77

14. Kwak N-S, Müller K-R, Lee S-W. A lower limb exoskeleton control system based on steady state visual evoked potentials. Journal of Neural Engineering. 2015 Aug 17; 12(5):056009. doi: 10.1088/1741-2560/12/5/056009 PMID: 26291321

15. Do AH, Wang PT, King CE, Chun SN, Nenadic Z. Brain-computer interface controlled robotic gait orthosis. Journal of NeuroEngineering and Rehabilitation. 2013 Dec 9; 10(1):1–9. doi: 10.1186/1743-0003-10-111

16. Kilicarslan A, Prasad S, Grossman RG, Contreras-Vidal JL. High accuracy decoding of user intentions using EEG to control a lower-body exoskeleton. In: International Conference of Engineering in Medicine and Biology Society (EMBC). 2013 Jul 3: 5606–5609.

17. Contreras-Vidal JL, Grossman RG. NeuroRex: A clinical neural interface roadmap for EEG-based brain machine interfaces to a lower body robotic exoskeleton. In: International Conference of Engineering in Medicine and Biology Society (EMBC). 2013 Jul 3: 1579–1582.

18. Cruz-Garza JG, Hernandez ZR, Nepaul S, Bradley KK, Contreras-Vidal JL. Neural decoding of expressive human movement from scalp electroencephalography (EEG). Front. Hum. Neurosci. 2014 Apr 8; 8 (188):1–16.

19. Ramos-Murguialday A, Broetz D, Rea M, Laer L, Yilmaz O, Brasil FL, et al. Brain-machine interface in chronic stroke rehabilitation: a controlled study. Annals of Neurology. 2013 Jul 1; 74(1):100–108. doi: 10.1002/ana.23879 PMID: 23494615

20. Morone G, Pisotta I, Pichiorri F, Kleih S, Paolucci S, Molinari M, et al. Proof of principle of a brain-computer interface approach to support poststroke arm rehabilitation in hospitalized patients: design, acceptability, and usability. Archives of Physical Medicine and Rehabilitation. 2015 Mar 31; 96(3):S71–S78. doi: 10.1016/j.apmr.2014.05.026 PMID: 25721550

21. Nilsson A, Vreede KS, Häglund V, Kawamoto H, Sankai Y, Borg J. Gait training early after stroke with a new exoskeleton—the hybrid assistive limb: a study of safety and feasibility. Journal of NeuroEngineering and Rehabilitation. 2014 Jun 2; 11(1):1–10. doi: 10.1186/1743-0003-11-92

22. Tsukahara A, Hasegawa Y, Eguchi K, Sankai Y. Restoration of gait for spinal cord injury patients using HAL with intention estimator for preferable swing speed. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2015 Mar; 23(2):308–318. doi: 10.1109/TNSRE.2014.2364618 PMID: 25350933

23. Venkatakrishnan A, Francisco GE, Contreras-Vidal JL. Applications of brain-machine interface systems in stroke recovery and rehabilitation. Current Physical Medicine and Rehabilitation Reports. 2014 Jun 1; 2(2):93–105. doi: 10.1007/s40141-014-0051-4 PMID: 25110624

24. Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller K-R. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Processing Magazine. 2008; 25(1):41–56. doi: 10.1109/MSP.2008.4408441

25. Suk H-I, Lee S-W. A novel Bayesian framework for discriminative feature extraction in brain-computer interfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2013 Feb; 35(2):286–99. doi: 10.1109/TPAMI.2012.69 PMID: 22431526

26. LaFleur K, Cassady K, Doud A, Shades K, Rogin E, He B. Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface. Journal of Neural Engineering. 2013 Aug 1; 10(4):046003. doi: 10.1088/1741-2560/10/4/046003 PMID: 23735712

27. Blankertz B, Lemm S, Treder M, Haufe S, Müller K-R. Single-trial analysis and classification of ERP components—a tutorial. NeuroImage. 2011 May 15; 56(2):814–25. doi: 10.1016/j.neuroimage.2010.06.048 PMID: 20600976

28. Yeom S-K, Fazli S, Müller K-R, Lee S-W. An efficient ERP-based brain-computer interface using random set presentation and face familiarity. PLoS ONE. 2014 Nov 10; 9(11):111157 doi: 10.1371/journal.pone.0111157

29. Müller-Putz GR, Scherer R, Brauneis C, Pfurtscheller G. Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. Journal of Neural Engineering. 2005 Dec 1; 2(4):123. doi: 10.1088/1741-2560/2/4/008 PMID: 16317236

30. Won D-O, Hwang H-J, Dähne S, Müller K-R, Lee S-W. Effect of higher frequency on the classification of steady-state visual evoked potentials. Journal of Neural Engineering. 2015 Dec 23; 13(1):016014. doi: 10.1088/1741-2560/13/1/016014 PMID: 26695712

31. Kim D-W, Hwang H-J, Lim J-H, Lee Y-H, Jung K-Y, Im C-H. Classification of selective attention to auditory stimuli: toward vision-free brain-computer interfacing. Journal of Neuroscience Methods. 2011 Apr 15; 197(1):180–185. doi: 10.1016/j.jneumeth.2011.02.007 PMID: 21335029

32. Hill NJ, Schölkopf B. An online brain-computer interface based on shifting attention to concurrent streams of auditory stimuli. Journal of Neural Engineering. 2012 Apr 1; 9(2):026011. doi: 10.1088/1741-2560/9/2/026011 PMID: 22333135

33. Müller-Putz GR, Scherer R, Neuper C, Pfurtscheller G. Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2006 Mar; 14(1):30–37. doi: 10.1109/TNSRE.2005.863842 PMID: 16562629

34. Breitwieser C, Kaiser V, Neuper C, Müller-Putz GR. Stability and distribution of steady-state somatosensory evoked potentials elicited by vibro-tactile stimulation. Medical and Biological Engineering and Computing. 2012 Apr 1; 50(4):347–357. doi: 10.1007/s11517-012-0877-9 PMID: 22399162

35. Chen X, Wang Y, Nakanishi M, Gao X, Jung TP, Gao S. High-speed spelling with a noninvasive brain-computer interface. Proceedings of the National Academy of Sciences. 2015 Nov 3; 112(44):E6058–E6067. doi: 10.1073/pnas.1508080112

36. Guger C, Allison BZ, Großwindhager B, Prückl R, Hintermüller C, Kapeller C, et al. How many people could use an SSVEP BCI? Frontiers in Neuroscience. 2012; 6:169. doi: 10.3389/fnins.2012.00169 PMID: 23181009

37. Lin Z, Zhang C, Wu W, Gao X. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. IEEE Transactions on Biomedical Engineering. 2006 Dec; 53(12):2610–2614. doi: 10.1109/TBME.2006.886577 PMID: 17152442

38. Zhang Y, Zhou G, Zhao Q, Onishi A, Jin J, Wang X, et al. Multiway canonical correlation analysis for frequency components recognition in SSVEP-based BCIs. In: Proceedings of the 18th International Conference on Neural Information Processing. Springer Berlin Heidelberg. 2011; 287: 295

39. Pan J, Gao X, Duan F, Yan Z, Gao S. Enhancing the classification accuracy of steady-state visual evoked potential-based brain-computer interfaces using phase constrained canonical correlation analysis. Journal of Neural Engineering. 2011 Jun 1; 8(3):036027. doi: 10.1088/1741-2560/8/3/036027 PMID: 21566275

40. Zhang YU, Zhou G, Jin J, Wang X, Cichocki A. Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis. International Journal of Neural Systems. 2014 Jun; 24(04):1450013. doi: 10.1142/S0129065714500130 PMID: 24694168

41. Luo A, Sullivan TJ. A user-friendly SSVEP-based brain-computer interface using a time-domain classifier. Journal of Neural Engineering. 2010 Apr 1; 7(2):026010. doi: 10.1088/1741-2560/7/2/026010

42. Zhang Y, Jin J, Qing X, Wang B, Wang X. LASSO based stimulus frequency recognition model for SSVEP BCIs. Biomedical Signal Processing and Control. 2012 Mar 31; 7(2):104–11. doi: 10.1016/j.bspc.2011.02.002

43. Zhang Y, Xu P, Cheng K, Yao D. Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface. Journal of neuroscience methods. 2014 Jan 15; 221:32–40. doi: 10.1016/j.jneumeth.2013.07.018 PMID: 23928153

44. Tello RMG, Müller SMT, Bastos-Filho T, Ferreira A. A comparison of techniques and technologies for SSVEP classification. In: 5th ISSNIP-IEEE Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC) 2014 May 26: 1–6.

45. Cecotti H, Graeser A. Convolutional neural network with embedded Fourier transform for EEG classification. In: Proceeding of the 19th International Conference on Pattern Recognition (ICPR). 2008 Dec 8: 1–4.

46. Cecotti H. A time-frequency convolutional neural network for the offline classification of steady-state visual evoked potential responses. Pattern Recognition Letters. 2011 Jun 1; 32(8):1145–53. doi: 10.1016/j.patrec.2011.02.022

47. Bevilacqua V, Tattoli G, Buongiorno D, Loconsole C, Leonardis D, Barsotti M, et al. A novel BCI-SSVEP based approach for control of walking in virtual environment using a convolutional neural network. In: International Joint Conference on Neural Networks (IJCNN). 2014 Jul 6: 4121–4128.

48. Lin YP, Wang Y, Jung TP. Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset. Journal of NeuroEngineering and Rehabilitation. 2014 Aug 9; 11(1):1–8. doi: 10.1186/1743-0003-11-119

49. McFarland DJ, Sarnacki WA, Vaughan TM, Wolpaw JR. Brain-computer interface (BCI) operation: signal and noise during early training sessions. Clinical Neurophysiology. 2005 Jan 31; 116(1):56–62. doi: 10.1016/j.clinph.2004.07.004 PMID: 15589184

50. Mikolajewska E, Mikolajewski D. Neuroprostheses for increasing disabled patients' mobility and control. Adv Clin Exp Med. 2012; 21(2):263–272. PMID: 23214292

51. Gwin JT, Gramann K, Makeig S, Ferris DP. Removal of movement artifact from high-density EEG recorded during walking and running. Journal of Neurophysiology. 2010 Jun 1; 103(6):3526–3534. doi: 10.1152/jn.00105.2010 PMID: 20410364

52. Kline JE, Huang HJ, Snyder KL, Ferris DP. Isolating gait-related movement artifacts in electroencephalography during human walking. Journal of Neural Engineering. 2015 Jun 17; 12(4):046022. doi: 10.1088/1741-2560/12/4/046022 PMID: 26083595

53. Bishop CM. Pattern recognition. Machine Learning. 2006; 128.

54. Montavon G, Braun ML, Müller K-R. Kernel analysis of deep networks. The Journal of Machine Learning Research. 2011 Feb 1; 12:2563–2581.

55. Cecotti H, Graser A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011 Mar; 33 (3):433–445. doi: 10.1109/TPAMI.2010.125 PMID: 20567055

56. Hecht-Nielsen R. Theory of the backpropagation neural network. In: International Joint Conference on Neural Networks (IJCNN). 1989 Jun 18; 1: 593–605.

57. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986 Oct; 323:533–536. doi: 10.1038/323533a0

58. LeCun YA, Bottou L, Orr GB, Müller K-R. Efficient backprop. In: Neural Networks: Tricks of the trade 2012 (pp. 9–48). Springer LNCS 7700 Berlin Heidelberg.

59. Lemm S, Dickhaus T, Blankertz B, Müller K-R. Introduction to machine learning for brain imaging. NeuroImage. 2011 May 15; 56(2):387–399. doi: 10.1016/j.neuroimage.2010.11.004 PMID: 21172442

60. Fazli S, Dähne S, Samek W, Bießmann F, Müller K-R. Learning from more than one data source: data fusion techniques for sensorimotor rhythm-based brain-computer interfaces. Proceedings of the IEEE. 2015 Jun; 103(6):891–906. doi: 10.1109/JPROC.2015.2413993

61. Dähne S, Bießmann F, Samek W, Haufe S, Goltz D, Gundlach C, et al. Multivariate machine learning methods for fusing functional multimodal. Proceedings of the IEEE. 2015; 103(9), 1507–1530. doi: 10.1109/JPROC.2015.2425807

62. Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS ONE. 2015 Jul 10; 10(7): e0130140. doi: 10.1371/journal.pone.0130140 PMID: 26161953

63. Samek W, Kawanabe M, Müller K-R. Divergence-based framework for common spatial patterns algorithms. IEEE Reviews in Biomedical Engineering. 2014; 7:50–72. doi: 10.1109/RBME.2013.2290621 PMID: 24240027

64. Contreras-Vidal JL. Identifying engineering, clinical and patient's metrics for evaluating and quantifying performance of brain-machine interface (BMI) systems. In: International Conference on Systems, Man and Cybernetics (SMC), 2014 Oct 5: 1489–1492.